Generative AI-Driven Semantic Communication Networks: Architecture, Technologies, and Applications

Abstract

Generative artificial intelligence (GAI) has emerged as a rapidly burgeoning field demonstrating significant potential in creating diverse content intelligently and automatically. To support such artificial intelligence-generated content (AIGC) services, future communication systems must fulfill stringent requirements, including high data rates, throughput, and low latency, while efficiently utilizing limited spectrum resources. Semantic communication (SemCom) has been deemed as a revolutionary communication scheme to tackle this challenge by conveying the meaning of messages instead of bit reproduction. GAI algorithms serve as the foundation for enabling intelligent and efficient SemCom systems in terms of model pre-training and fine-tuning, knowledge base construction, and resource allocation. Conversely, SemCom can provide AIGC services with low latency and high reliability due to its ability to perform semantic-aware encoding and compression of data, as well as knowledge- and context-based reasoning. In this survey, we break new ground by investigating the architecture, wireless communication schemes, and network management of GAI-driven SemCom networks. We first introduce a novel architecture for GAI-driven SemCom networks, comprising the data plane, physical infrastructure, and network control plane. In turn, we provide an in-depth analysis of the transceiver design and semantic effectiveness calculation of end-to-end GAI-driven SemCom systems. Subsequently, we present innovative generation level and knowledge management strategies in the proposed networks, including knowledge construction, update, and sharing, ensuring accurate and timely knowledge-based reasoning. Finally, we explore several promising use cases, i.e., autonomous driving, smart cities, and the Metaverse, to provide a comprehensive understanding and future direction of GAI-driven SemCom networks.